1-1-2012

Assessing Metacognitive Skills Using Adaptive Neural Networks

Anderson Justin

Kouider Mokhtari

Arun Kulkarni
University of Texas at Tyler, akulkarni@uttyler.edu

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 12 (2012) 294 – 299

Procedia
Computer Science

Complex Adaptive Systems, Publication 2
Cihan H. Dagli, Editor in Chief
Conference Organized by Missouri University of Science and Technology
2012- Washington D.C.

Assessing Metacognitive Skills Using Adaptive Neural Networks

Justin Anderson, Kouider Mokhtari, Arun Kulkarni

The University of Texas at Tyler, Tyler, TX 75799, USA

Abstract

The assessment of students' levels of metacognitive knowledge and skills is critical in determining their ability to effectively perform complex cognitive tasks such as solving mathematics or reading comprehension problems. In this paper, we use an adaptive multiplayer perceptron model to categorize participants based on their metacognitive awareness and perceived use of reading strategies while reading. Eight hundred and sixty-five middle school students participated in the study. All participants completed a 30-item instrument- the Metacognitive Awareness-of-Reading Strategies Inventory (MARSI). We used adaptive multi-layer perceptron models to classify participants into three groups based on their metacognitive strategy awareness levels using thirteen and nine attributes representing problem-solving and support reading strategies. The architecture for the neural network models is based on the input data. The number of units in the input layer is equal to the number of attributes and the number of units in the output layer is equal to the number of categories. We classified participants into three categories based on the level of awareness. The models are evaluated using the measures such as user's efficiency and Kappa coefficient that are obtained from the error matrix. We obtained an overall efficiency of 86.92 and 81.89 percent with 13 and 9 input features, respectively. The results indicate that once the network is trained, it can be used to assess students' metacognitive awareness and use of reading strategies with the help of observed attributes.

Keywords: Cognitive Reading Skills; Classification; Mutli-layer Perceptron

1. Introduction

The assessment of students' levels of metacognitive knowledge and skills is critical when determining their ability to effectively perform complex cognitive tasks such as solving mathematics or reading comprehension problems. For all intents and purposes, reading comprehension is a metacognitive or problem-solving task. Knowledge of cognition involves (a) reader variables (e.g., knowledge of ones' abilities, strengths and weaknesses with respect to reading), (b) task variables (e.g., knowledge of reading tasks involved, level of difficulty involved, as well as mental or tangible resources necessary to accomplish the tasks at hand), and (c) procedural knowledge variables (e.g., knowledge of which strategies to use, as well as knowledge of how, when, and why to use these strategies).

Research within the domains of cognition and reading comprehension has led to an increasing emphasis on the role of metacognitive awareness and control of one's thinking processes while engaged in cognitive tasks such as reading comprehension. This research has contributed a great deal to our understanding of students' cognitive and

 doi:10.1016/j.procs.2012.09.072

metacognitive strategy development. It has also shed a new light on the influence in classrooms of student perceptions of their metacognitive knowledge, cognitive skills, and reading strategies, which are generally defined as systematic cognitive plans that assist in the acquisition of information and task performance (Flavell, 1979; Pressley, 2000).

There is agreement among researchers that variability in reader characteristics can be used to partially explain differences in reading comprehension performance. The process of reading is greatly influenced by the beliefs, attitudes, and values that readers possess. We know from research that what learners feel and know about their own abilities and skills may affect whether they succeed or fail in school (Paris & Winograd, 1990). Indeed, the development of metacognitive beliefs about reading and the understanding of the parameters and complexities involved in reading tend to develop whenever and wherever students receive instruction in reading.

Individual differences in metacognitive awareness have been found among students varying in ability levels. Research has shown that high ability readers tend to exhibit higher levels of metacognitive awareness about reading processes than do low ability readers. For example, in a recent study aimed at examining differences in the metacognitive awareness of reading strategies among 350 United States (US) and English as a Second language (ESL) university students, Sheorey & Mokhtari (2001) found that both US and ESL students demonstrated a high level of awareness of nearly 30 reading strategies. They also found that US female students reported a significantly higher usage of reading strategies than did their male counterparts, and that the use of reading strategies was associated with higher levels of reading ability for both groups of students. These findings are consistent with other studies (e.g., Jimenez, Garcia, & Pearson, 1995) that have shown that efficient bilingual readers exhibit an awareness of a rich supply of strategies when reading in English and Spanish, and that they do make use of such when reading English.

While popular statistical techniques such as regression have often been used to predict and diagnose students' reading difficulties, there is an increasing interest among educational researchers and practitioners to use Artificial Neural Networks (ANN) as useful tools for quick, reliable, and flexible identification of reading difficulties among struggling readers. First, to our knowledge, the use of ANNs in educational research is relatively new and largely unexplored as a prospective method for predicting and diagnosing reading problems. Second, when considering the idiosyncrasies of the data used in our study, which seeks to uncover relationships among student variable such as age, gender, grade level, ethnicity, perceptions of self as a reader in relation to their levels of metacognitive strategy awareness and perceived use of reading strategies, we find that ANNs provide a set of suitable data-driven systems that help detect complex patterns within multi-dimensional data, and they do not necessarily depend on assumptions of functional form, probability distribution, or smoothness. Third, we see potential benefits of using Multi-Layer Perceptron (MLP) within ANNs for educational applications such as prediction and diagnosis of reading difficulties, in part due to their ability to use non-linear transformations and to learn patterns of behavior among independent (or input) and dependent (or output) variables.

2. Method

In the present study, we applied neural networks to identify levels of metacognitive awareness and use of reading strategies among middle school students. The underlying goal was to explore the relationships that might exist between students' perceptions of their own metacognition, and reading comprehension performance. Such information will help a great deal in diagnosing students' awareness of their own reading processes, documenting their actual usage of reading strategies while reading, and designing instructional programs aimed at helping these students become strategic, thoughtful, and constructively responsive readers (Pressley & Afflerbach, 1995).

2.1 Data Collection

For purposes of this study, we collected two sets of data pertaining to the participants' (a) demographics (e.g., age, grade level, gender, ethnicity, and perceptions of their ability to read), and (b) perceived awareness and use of reading strategies, which are organized in three categories, namely Global, problem-solving, and support reading strategies. All participants completed a 30-item instrument—the *Metacognitive-Awareness-of-Reading-Strategies Inventory* [MARSI] (Mokhtari & Reichard, 2002), which was specifically designed for measuring students' metacognitive awareness and use of reading strategies while reading academic or school-assigned materials.

The MARSI instrument measures three broad categories of strategies including (1) Global Reading Strategies (GLOB) which can be thought of as generalized, or global reading strategies aimed at setting the stage for the reading act (e.g., setting purpose for reading, previewing text content, predicting what the text is about, etc.); (2) Problem-Solving Strategies (PROB) which are localized, focused problem-solving or repair strategies used when problems develop in understanding textual information (e.g., checking one's understanding upon encountering conflicting information, re-reading for better understanding, etc.); and (3) Support Reading Strategies (SUP) which provide the support mechanisms or tools aimed at sustaining responsiveness to reading (e.g., use of reference materials like dictionaries and other support systems). These three classes of strategies interact with and support each other when used in the process of constructing meaning from text.

The data analyzed for purposes of this study is a part of a larger study designed to determine the ways in which students' awareness of reading strategies contributed to their overall reading comprehension performance. The data, which we collected in the spring 2010, consisted of (a) basic demographic information such as age, gender, ethnicity, and perceptions of reading ability, and (b) measures of students' awareness of reading strategies. Because our primary goal was to apply neural networks to identify levels of metacognitive awareness and use of reading strategies among middle school students, we considered only one of the three types of metacognitive awareness of reading strategies, namely Problem Solving Strategies. Levels of awareness and use of problem solving strategies considered included low, medium, and high levels of strategy awareness and use. The inputs or variables collected from our sample of 856 students are summarized in Table 1.

Table 1. Description of Variables Used in the Analyses

Variable	**Description**
Age	Student age (Mean age)
Grade	Student current grade placement (6^{th}, 7^{th}, & 8^{th} grade)
Ethnicity	Student ethnic membership (e.g., Caucasian, Hispanic)
Gender	Student gender membership (e.g., male or female)
Reader Perception	Student perception of own reading ability (1=Excellent; 4=Poor)
MARSI Problem-Solving Strategies (n=08)	Eight reading strategies related to student awareness

2.2 Data Analysis

We used a neural network to analyze the data set. The model is shown in Figure 1. It is a three-layer network with a back-propagation learning algorithm. Layer L_1 is the input layer. Layer L_2 is the hidden layer, and L_3 is the output layer. The number of units in the input layer is equal to the number of features. Layer L_3 consists of a single unit. Units in layers L_2 and L_3 employ the Sigmoid activation function. The network represents a three layer feed-forward network with a back-propagation learning algorithm. The output consists of three units that represent three categories of problem solving skills-high, medium and low. The network has two phases: the learning phase and the decision making phase. During the decision-making phase, it maps the input vector. The output vectors {1, 0, 0}, {0, 1, 0}, and {0, 0, 1} represent three categories low, medium, and high, respectively. The learning algorithm is described below (Kulkarni, 2001).

Step 1: Initialize the weights. The weights between layers L_1L_2 and L_2L_3 are represented by elements of matrices P and Q. These weights are initialized to small random values so that the network is not saturated by large values of weights. Let n and m, represent the number of units in layers L_1, and L_2, respectively. Let l represent units in L_3. In this case l is equal to 1

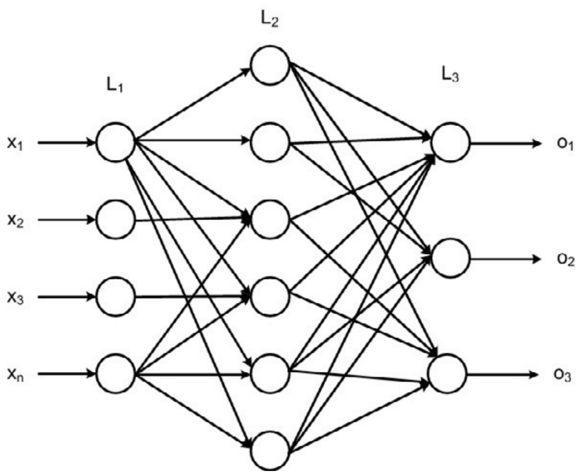

Figure 1. Three layer feed-forward network

.

Step 2: Present a continuous-valued input vector $\mathbf{x} = \left(x_1, x_2, \ldots, x_n\right)^T$ to layer L_1 and obtain the output vector $\mathbf{o} = \left(o_1, o_2 \ldots o_l\right)$ at layer L_3. In order to obtain the output at L_3 calculation is done layer by layer starting from layer L_1 to L_3 using Equations (1) and (2). In these equations $net_i, x_j,$ *and* w_{ij} represent net input for unit *i*, input *j*, and the weight, respectively.

$$net_i = \sum_{j=1}^{n} x_j w_{ij} \tag{1}$$

$$o_i = \frac{1}{1 + \exp\left(-net_i\right)} \tag{2}$$

Step 3: Calculate change in weights. In order to do this the output vector **o** is compared with the desired output vector or the target vector **d**, and the error between the two vectors is obtained. The error is then propagated backward to obtain the change in weights Δq_{ij} that is used to update the weights. Δq_{ij} for weights between layers L_2L_3 is given by:

$$\Delta q_{ij} = -\alpha \frac{\partial E}{\partial q_{ij}} \tag{3}$$

Equation (3) can be reduced to

$$\Delta q_{ij} = \alpha \delta_i o_j \tag{4}$$

where α is a training rate coefficient (typically 0.01 to 1.0), o_j is the output of neuron *j* in layer L_2, and δ_i is given by

$$\delta_i = \left(d_i - o_i\right) o_i \left(1 - o_i\right) \tag{5}$$

In Equation (5), o_i represents the actual output of neuron *i* in layer L_3, and d_i represents the target or the desired output at neuron *i* in layer L_3. The back-propagation algorithm trains the hidden layers by propagating the output error back through layer by layer, adjusting weights at each layer. The change in weights between layers L_1L_2 can be obtained as

$$\Delta p_{ij} = -\beta o_j \delta_{Hi} \tag{6}$$

where β is a training rate coefficient for layer L_2 (typically 0.01 to 1.0), o_j is the output of neuron *j* in layer L_1, and

$$\delta_{Hi} = o_i\left(1-o_i\right)\sum_{k=1}^{m}\delta_k q_{ik} \tag{7}$$

In Equation (7), o_i is the output of neuron *i* in layer L_2, and the summation term represents the weighted sum of all δ_k values corresponding to neurons in layer L_3 that are obtained by using Equation (5)

Step 4: Update the weights.

$$\begin{aligned} q_{ij}\left(k+1\right) &= q_{ij}\left(k\right)+\Delta q_{ij} \\ p_{ij}\left(k+1\right) &= p_{ij}\left(k\right)+\Delta p_{ij} \end{aligned} \tag{8}$$

Where $q_{ij}\left(k+1\right)$ and $p_{ij}\left(k+1\right)$ represent values of the weights at iteration $k+1$ (after adjustment), and $q_{ij}\left(k\right)$ *and* $p_{ij}\left(k\right)$ represent the values of the weights at iteration *k*.

Step 5: Obtain the mean squared error ε for neurons in layer L_3.

$$\varepsilon = \frac{1}{2}\sum_{i=1}^{n}\left(o_i - d_i\right)^2 \tag{9}$$

If the error ε is greater than some minimum $\varepsilon_{\min}$, then repeat steps 2 through 4; otherwise terminate the training process. The learning algorithm for the first model is same as the second model. However, the first model does not have the hidden layer; the change in weights is given by Equation (5).

3. Results and Discussion

We developed software to simulate the neural network model, and analyzed the data set using a selected number of features. The simulator determines the structure of the network based on the number of input attributes and output classes. We analyzed data using three models. In the first model we used 13 features that included 8 features from problem solving strategies and the first 5 features as shown in Table 1. We used the mean vector obtained from randomly selected half samples to train the network. The remaining half samples were used as test samples. We obtained the overall efficiency of 86.92% with the first model. We analyzed data with the second model with the reduced number of features. In the second model we used 9 features that included 4 features from problem solving strategies and the first 5 features as shown in Table 1. We obtained the overall efficiency of 81.89% with the second model. In the second model, we used features that showed the maximum variance. The results demonstrate that once we train a network it can be successfully used to classify any student record into a category such as high, medium, or low level of awareness. Based on this study, we believe a MLP network can be used as a promising tool for predicting and diagnosing students' levels of metacognitive awareness and use of reading strategies on the basis of multiple variables related to reader characteristics.

References

J. H. Flavell. Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist*, 34, 10, (1979). 906-911.

R. Jimenez, G. Garcia, and P. Pearson. Three children, two languages, and strategic reading: Case studies in bilingual/monolingual reading. *American Educational Research Journal,* 32, 1, (1995). 67-97

K. Mokhtari and C. Reichard. Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology,* 94, 2, (2002). 249-259.

A. Kulkarni. *Computer vision and fuzzy neural systems*. Prentice Hall PTR, Upper Saddle River, New Jersey. (2001). 108-128

S. G. Paris, and P. Winograd. How metacognition can promote academic learning and instruction. In B. F. Jones & L. Idol (Eds.), *Dimensions of Thinking and Cognitive Instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates. (1990).

M. Pressley. What should comprehension instruction be the instruction of? In Kamil, M., Mosenthal, P., Pearson, P.D., & Barr, R. (Eds.), *Handbook of Reading Research* (Vol. III), Mahwah, NJ: Lawrence Erlbaum. (2000). 545-561.

M. Pressley, and P. Afflerbach. *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, New Jersey: Lawrence Erlbaum Associates (1995).

R. Sheorey, & K. Mokhtari. Coping with academic materials: Differences in the reading strategies of native and non-native readers. *System: An International Journal of Educational Technology and Applied Linguistics,* 29, 4, (2001). 431-449.